Fusing Directions and Displacements in Translation Averaging
Supplementary Material

Lalit Manam and Venu Madhav Govindu
Indian Institute of Science
Bengaluru, India - 560012
{lalitmanam,venug}@iisc.ac.in

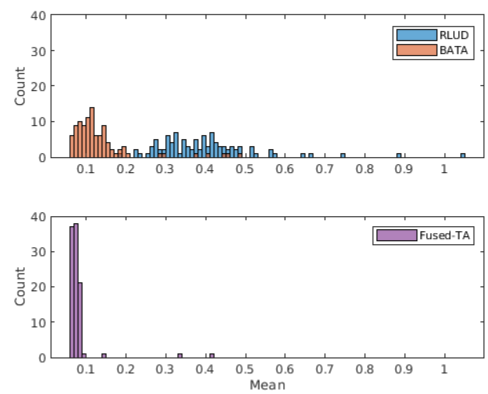
A. Implementation Details

In Sec. 4.2 of the main paper, we provided Algo. 1 to fuse the solutions obtained from the relaxed displacement and direction-based costs iteratively. Here, we provide the implementation details for Algo. 1. The translations are initialized with RLUD, as done in [45]. We use Cauchy robust loss for e_{rdis} and e_{rdir} with its scale factor $\alpha = 0.1$. Algo. 1 is run until the relative change in both the costs is lesser than 10^{-6} and the absolute change in the translations is lesser than 10^{-5}. The loop is run for a maximum of 100 iterations.

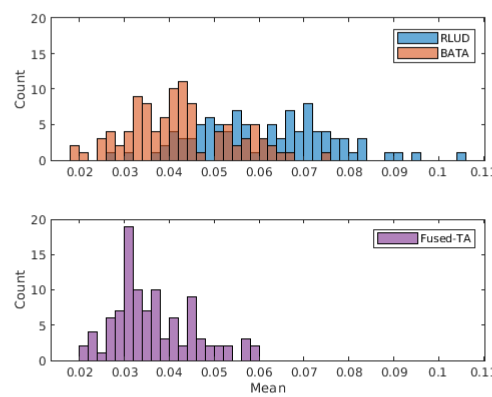
B. Additional Results on Synthetic Data

In Sec. 5.1 of the main paper, we presented results for the high noise case, i.e. $\sigma = 5$ in Fig. 3 (of the main paper). Here, we present the results for $\sigma = 2$ for both datasets having disparate (Syn_{Diff}) and similar baselines (Syn_{Sim}).

Fig. S1 shows the histogram of mean errors obtained from 100 runs for noise level $\sigma = 2$ on both datasets. It can be seen that, for both the datasets Syn_{Diff} and Syn_{Sim} with $\sigma = 2$ noise, the BATA has lower mean errors compared to RLUD in most of the instances, which was not the case for Syn_{Sim} with $\sigma = 5$ noise (Fig. 3b of the main paper). This reveals that the performance of the relaxed costs e_{rdis} (used in RLUD) and e_{rdir} (used in BATA) is dependent on both the spread of the cameras and the noise level in the input directions. Fused-TA performs better for both baseline conditions compared to both RLUD and BATA, with more instances having the least mean errors. This shows the advantage of fusing the solutions of both the relaxed costs iteratively, leading to better translation estimates.



(a) $Syn_{Diff}, \sigma = 2$

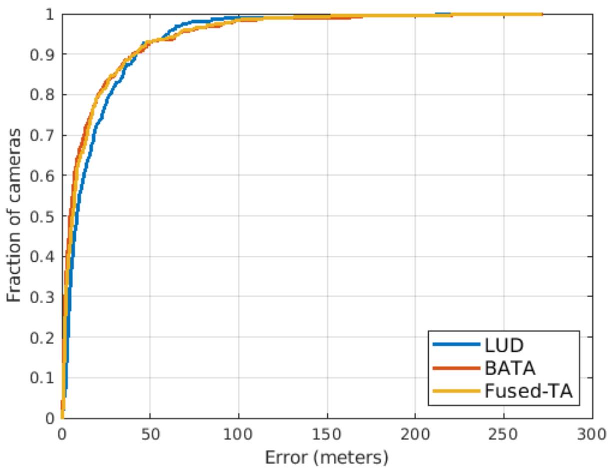


(b) $Syn_{Sim}, \sigma = 2$

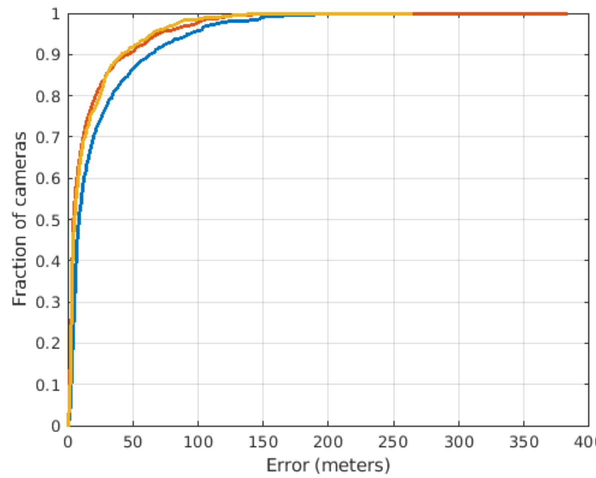
Figure S1. Histogram of mean errors for the two synthetic datasets with $\sigma = 2$ noise. The leftward shift indicates superior performance of our method.

C. Additional Results on Real Data

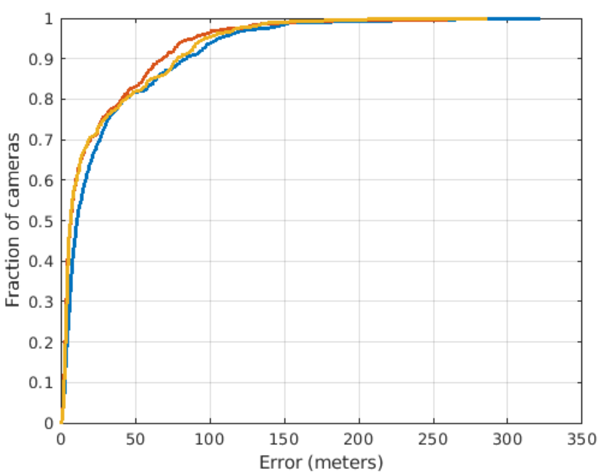
In Fig. 4 of the main paper, we presented a zoomed part of the empirical cumulative error distribution of the camera translations obtained using different methods. Here, we



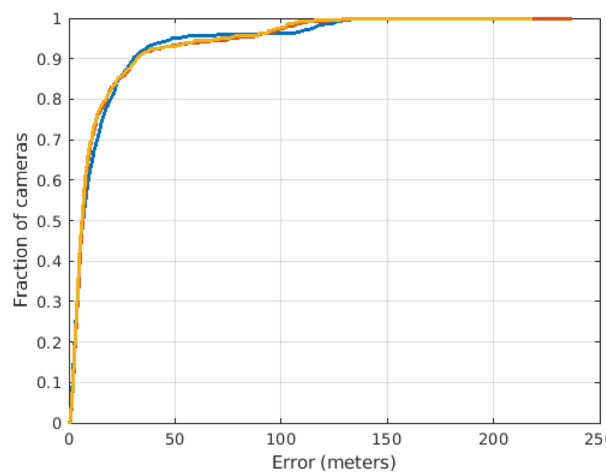
(a) Madrid Metropolis



(b) Roman Forum



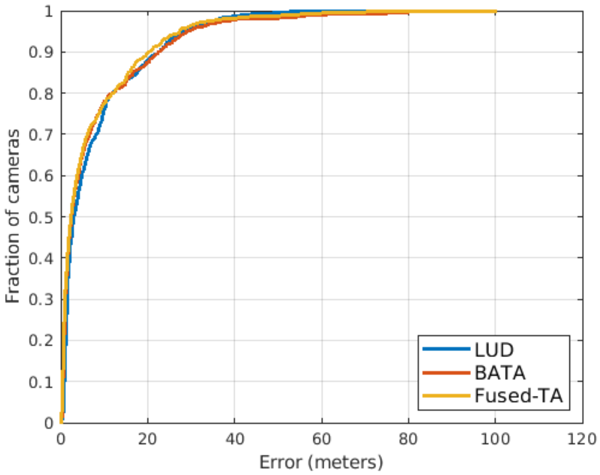
(c) Tower of London



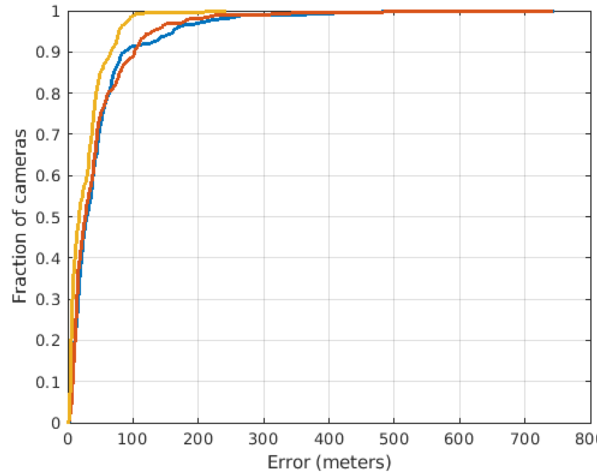
(d) Vienna Cathedral

Figure S2. Empirical cumulative error distribution (in meters) for camera translations obtained on 1DSfM datasets.

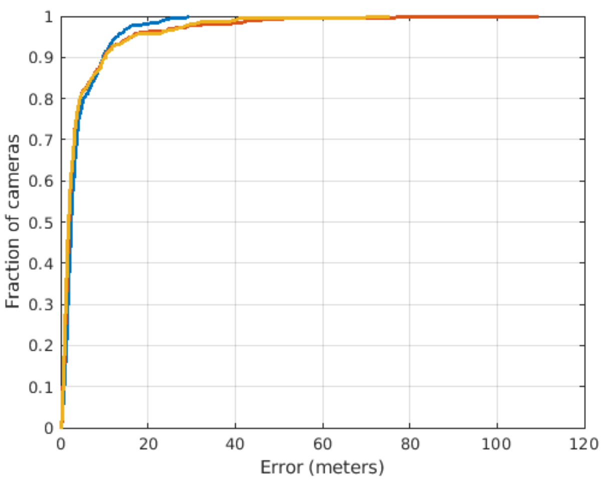
provide the complete plots in Fig. S2. As can be seen in the plots, for some cameras, the performance of LUD is better than that of BATA, and for the other cameras, it is vice versa. In the case of Fused-TA (our method), the performance of the cameras with low errors ($\leq$ 50 meters) is either best or close to the best-performing method among LUD or BATA. But for cameras with high errors ($>$ 50 meters), the performance of our method lies between LUD and BATA. This shows that Fused-TA tries to incorporate the benefits from both costs. In Fig. S3, we provide empirical cumulative distribution errors on other 1DSfM datasets. In these datasets, we again see that some cameras are recovered better in LUD while other cameras are recovered better with BATA. Fused-TA performs the best for cameras with low errors, and its performance lies in between LUD and BATA for cameras with high errors.



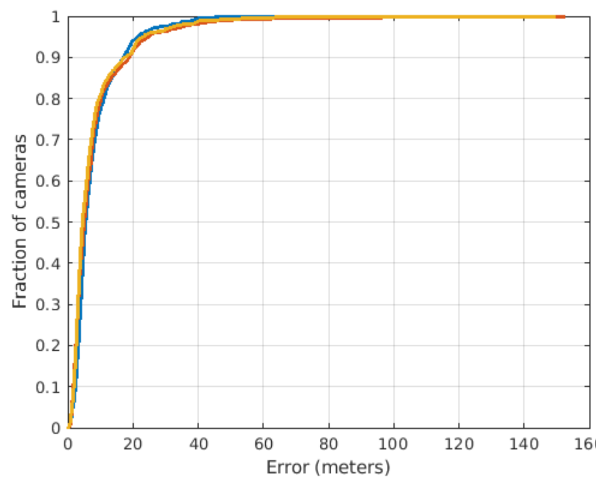
(a) Alamo



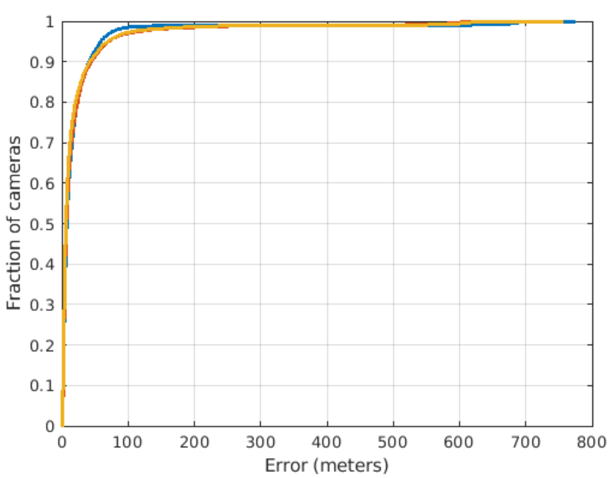
(b) Ellis Island



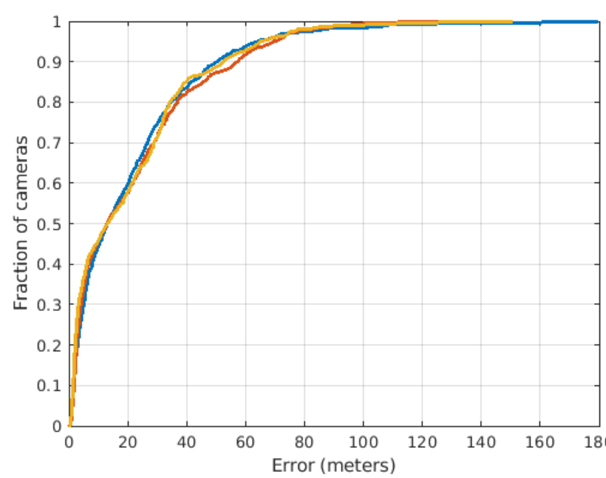
(c) Montreal Notre Dame



(d) Piazza del Popolo



(e) Trafalgar



(f) Yorkminster

Figure S3. Empirical cumulative error distribution (in meters) for camera translations obtained on a few other 1DSfM datasets.